CERTLABZ CERTIFICATION BOARD

Certified Computer Vision Engineer (CCVE)

Ten Sample Questions

These ten questions are written at the difficulty of the live exam. Six carry a figure, a data table, a configuration or a code listing; four are plain scenario questions, as on the live paper. Every one is a scenario a working computer vision engineer would meet, and none of them appears in the live question bank.

The live exam is multistage adaptive: how you score on one stage selects the difficulty of the next, so no two candidates sit the same paper. It is a performance-based exam: each form can carry up to 5 to 12 performance-based questions alongside the multiple-choice questions, which means you build, sequence or complete something rather than only choosing from a list. Answers and full explanations follow each question here; the live exam does not disclose item-level answers.

Question 1 — Domain 3: Detection, Segmentation and Tasks — DATA

Emma Wilson matches vision tasks to their outputs.

Task	Output
Classification	One label for the whole image
Object detection	Boxes and classes for objects
Semantic segmentation	A class for every pixel
Instance segmentation	Separate masks per object

Table 1.1 Tasks and outputs.

A system must count and separately follow each pedestrian in a video. Which task fits best?

A. Whole-image classification, because a single label per frame is enough to count and follow each pedestrian in the scene over time.

B. Detection with tracking (or instance segmentation), because it separates each pedestrian and follows their identity across the video frames.

C. Semantic segmentation alone, because labelling every pixel as person or not is sufficient to count and follow each individual separately.

D. Optical character recognition, because reading any text in the scene is the most reliable way to count and track the pedestrians present.

Correct answer: B
Counting and separately following each pedestrian requires instance-level output plus temporal continuity, so detection with tracking (or instance segmentation) fits. Whole-image classification gives no locations, semantic segmentation does not separate individuals, and OCR reads text rather than tracking people.

Question 2 — Domain 4: Data, Training and Evaluation — DATA

James Carter evaluates a defect detector where defects are 1% of images.

Metric	Value
Overall accuracy	99%
Defects in the data	1%
Defects actually caught (recall)	10%
Behaviour	Predicts "no defect" most of the time

Table 2.1 Detector results.

Why is the 99% accuracy misleading, and what should be used instead?

A. It is not misleading; 99% accuracy proves the detector is excellent and no other metric needs to be considered for this problem.

B. Class imbalance makes accuracy deceptive; predicting "no defect" scores 99% while missing defects, so use recall and precision instead.

C. The accuracy is too low, and the fix is simply to train the model for more epochs until the overall accuracy figure reaches 100%.

D. The metric is fine but the threshold is wrong, and lowering the classification threshold alone will make the 99% accuracy meaningful.

Correct answer: B
With defects at 1%, a model that almost always predicts "no defect" reaches 99% accuracy while catching only 10% of real defects. Accuracy is deceptive under class imbalance, so precision and recall (and metrics like F1 or mAP) reveal the true performance. More epochs or a threshold tweak does not fix the wrong choice of metric.

Question 3 — Domain 5: Deployment and Optimization — CONFIG

Olivia Davis must deploy a detector to a camera that needs 30 fps but the model runs at 5 fps.

```
target : 30 fps on the device
current: 5 fps with the full model
options: quantise, prune, distil, smaller architecture,
         hardware acceleration (NPU)
```

Listing 3.1 The deployment constraint.

What is the right approach?

A. Accept 5 fps as adequate, because a more accurate large model at a low frame rate is always preferable to a faster optimised one.

B. Optimise the model (quantise, prune or distil, or use a smaller architecture) and use hardware acceleration to reach the frame-rate target.

C. Increase the input image resolution, because higher-resolution frames will make the existing model process each frame faster on the device.

D. Move all inference to a distant cloud region, because cloud processing is always faster than any optimised model running on the edge device.

Correct answer: B
Meeting a real-time budget on-device means optimising the model with quantisation, pruning, distillation or a smaller architecture, and using hardware acceleration, validating the accuracy trade-off. Accepting 5 fps misses the requirement, raising resolution makes it slower, and a distant cloud adds latency that defeats real-time edge use.

Question 4

Domain 2: Deep Learning for Vision

DATA

Liam Anderson has 800 labelled images and a large CNN that overfits when trained from scratch.

Option	Approach
A	Train an even larger CNN from scratch
B	Use transfer learning plus data augmentation
C	Remove augmentation to simplify
D	Reduce image resolution only

Table 4.1 Options for the small dataset.

Which option is most likely to work, and why?

A. Option A, because a larger model trained from scratch will always overcome a small dataset given enough parameters and training time.

B. Option B, because a pretrained model already learned general features and augmentation expands the effective data, so few labels are needed.

C. Option C, because removing augmentation gives the model cleaner data and is the most reliable cure for overfitting on a small dataset.

D. Option D, because lowering the resolution alone reduces overfitting and is sufficient to make a large CNN generalise from 800 images.

Correct answer: B
With only 800 images, transfer learning leverages features already learned on a large dataset, and augmentation expands the effective training data, which together address overfitting. A larger from-scratch model overfits more, removing augmentation reduces data variety, and lowering resolution alone does not solve the shortage of labels.

Question 5

Domain 1: Image Fundamentals and Classical Vision

CONFIG

Ava Thompson finds an edge model works in the lab but fails on deployed cameras.

```
training preprocessing: resize 224x224, normalise mean/std
device preprocessing  : resize 256x256, no normalisation
symptom               : accuracy far lower on the cameras
```

Listing 5.1 Preprocessing in each environment.

What is the cause, and what is the fix?

A. The model is too small for the cameras, so training a larger model will fix the accuracy drop without changing any preprocessing steps.

B. Preprocessing differs between training and the device, so the model sees a shifted distribution; align device preprocessing with training.

C. The cameras are simply lower quality, so the only solution is to replace all of the deployed cameras with much higher-resolution units.

D. The accuracy figure in the lab was wrong, so re-reporting the lab metric to match the device result resolves the discrepancy fully.

Correct answer: B
The device resizes to a different size and skips normalisation, so the model receives inputs from a distribution it was not trained on, a preprocessing skew that degrades accuracy. Aligning device preprocessing to match training fixes it. A bigger model, new cameras, or re-reporting metrics does not address the mismatch.

Question 6

Domain 6: Ethics, Robustness and Production

DATA

A facial analysis system is measured across demographic groups before deployment.

Group	Accuracy
Group A	96%
Group B	95%
Group C	74%
Group D	72%

Table 6.1 Accuracy by group.

What must the team do before deploying this high-impact system?

A. Deploy immediately, because the two strongest groups show high accuracy and the overall average across groups is acceptable for launch.

B. Address the disparity with better data and mitigation, evaluate per group, and for a high-impact use reconsider or add safeguards before deploying.

C. Report only the overall average accuracy, because presenting a single blended number is the clearest and most honest way to describe the system.

D. Remove groups C and D from the evaluation, because dropping the lower-scoring groups makes the reported results consistent and ready to ship.

Correct answer: B
The system performs far worse for groups C and D, a serious fairness problem for a high-impact use. The team must test per group, improve the data and mitigate the disparity, and for such a consequential system reconsider deployment or add safeguards. Deploying anyway, hiding behind an average, or dropping the weak groups would cause and conceal real harm.

Question 7 Domain 2: Deep Learning for Vision

A model trained only on daytime images performs poorly at night. What is the issue, and the right response?

A. The model is simply too small, so training a larger network on the same daytime-only images will make it perform well at night too.

B. Domain shift: the model never saw night conditions, so include representative night-time data and augmentation and test on night images.

C. Night images cannot be learned by any vision model, so the only option is to restrict the system to daytime operation permanently.

D. The GPU is too slow at night, so upgrading the inference hardware will restore the model's accuracy on the night-time images.

Correct answer: B
Poor night performance is domain shift: the training data did not cover night conditions, so the model faces a distribution it never learned. The fix is to include representative night-time data and augmentation and to test on night images. Model size and hardware do not address missing coverage, and night imagery is certainly learnable.

Question 8 Domain 4: Data, Training and Evaluation

A model scores very well on the test set, but the test images came from the same recording session as the training images. Why is this a problem?

A. It is not a problem, because using test images that are very similar to the training images gives the most reliable performance estimate.

B. The test data is not truly independent, so the score is inflated; the test set must be genuinely held out and representative of real use.

C. The only issue is that the test set is too small, and simply adding more images from the same session would make the estimate trustworthy.

D. The problem is that the model is too accurate, so deliberately weakening it until the test score falls would make the evaluation valid.

Correct answer: B
When test images share a session (and thus subjects, lighting and background) with training, the test set is not independent and the reported score is optimistic. An honest estimate needs a truly held-out test set that reflects real deployment conditions. More images from the same session, or weakening the model, does not fix the lack of independence.

Question 9 Domain 5: Deployment and Optimization

After quantising a vision model to run on a device, accuracy drops more than expected. What is a reasonable response?

A. Ship the quantised model as-is, because any model that runs on the target device is preferable to one that is more accurate but slower.

B. Use quantisation-aware training or a less aggressive scheme, then re-evaluate the accuracy-versus-efficiency balance for the task.

C. Abandon on-device deployment entirely, because a noticeable accuracy drop after quantisation proves the model cannot run on the edge at all.

D. Increase the input resolution to compensate, because higher-resolution inputs always restore the accuracy lost during model quantisation.

Correct answer: B

An unexpected accuracy drop from quantisation is addressed by quantisation-aware training or a less aggressive scheme, then re-checking the accuracy-efficiency balance for the task. Shipping a badly degraded model, abandoning the edge, or simply raising resolution are not sound responses to the trade-off.

Question 10 Domain 6: Ethics, Robustness and Production

A road-sign detector can be fooled by a small printed sticker placed on a stop sign. What concern is this, and what helps address it?

A. It is harmless, because a sticker is a physical object and cannot possibly change how a computer vision model classifies a road sign.

B. It is an adversarial and robustness risk; test against such attacks and design defences and fail-safes, especially for safety-critical use.

C. It is only a data-labelling error, so relabelling the stop-sign images in the training set is all that is needed to remove the vulnerability.

D. It is purely a resolution problem, so running the detector on higher-resolution frames alone will guarantee the sticker can never fool it.

Correct answer: B

A sticker that causes misclassification is a physical-world adversarial attack, a serious robustness and safety concern for a system like this. Addressing it means testing against such attacks and designing defences and fail-safes. It is not harmless, not merely a labelling or resolution issue, and safety-critical vision demands this scrutiny.

Published by the CertLabz Certification Board, a division of CertLabz LLC. Companion to objectives document CCVE-001. https://certlabz.com/certifications/certified-computer-vision-engineer.html